IOWA STATE UNIVERSITY
Digital Repository

Graduate Theses and Dissertations

Iowa State University Capstones, Theses and Dissertations

2019

Creating probabilistic streamflow forecasts using HRRRE & HREF probabilistic quantitative precipitation forecasts

Andrew Goenner
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

 Part of the Hydrology Commons, and the Meteorology Commons

Recommended Citation
Goenner, Andrew, "Creating probabilistic streamflow forecasts using HRRRE & HREF probabilistic quantitative precipitation forecasts" (2019). *Graduate Theses and Dissertations*. 17017.
https://lib.dr.iastate.edu/etd/17017

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

Creating probabilistic streamflow forecasts using HRRRE & HREF probabilistic quantitative precipitation forecasts

by

Andrew Raymond Goenner

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Meteorology

Program of Study Committee:
Kristie J. Franz, Co-major Professor
William A. Gallus, Co-major Professor
Mark S. Kaiser

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this thesis. The Graduate College will ensure this thesis is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2019

Copyright © Andrew Raymond Goenner, 2019. All rights reserved.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

First, I want to thank Jehovah Jireh for giving me the abilities and opportunities to learn about His creation and therefore Himself. Next, I would like to thank Dr. William Gallus and Dr. Kristie Franz for taking me on as their student, leading and guiding me through the experience of learning thereby deepening of my education. Furthermore, I would like to thank NOAA and Iowa State University for providing the funding that supported my research during my time here. Also, I would like to thank my family for guiding, teaching and helping me throughout my childhood, supporting in my endeavors, through today. Additionally, thank you to other professors, Dr. Kaiser, Dr. Zhou, Dr. Wu, and Dr. Gutowski, who helped in teaching or directing me while at Iowa State University. Finally, I would like to thank my fellow classmates and friends who helped to answer problems and making this an enjoyable experience for the past two years.

ABSTRACT

The present study examines how skillful probabilistic streamflow forecasts are when using convection-allowing ensemble models' probabilities of precipitation exceeding specified threshold accumulations as input. Both the High-Resolution Rapid Refresh Ensemble (HRRRE) and High-Resolution Ensemble Forecast version 2.0 (HREF) output were tested. A vital component of this work was the creation of expected rainfall amounts at every grid point for seven different probability of exceedance values. The rainfall amounts for each of the probability of exceedance values were calculated using cubic interpolation from the probabilistic quantitative precipitation forecasts (PQPFs) generated from the HRRRE and the HREF models by use of a Gaussian smoothing technique applied by the HRRRE and HREF developers. The grid point precipitation amounts associated with the probability of exceedance values were then inputted into a hydrologic model for 11 different river basins across the upper Midwest for 109 cases during June, July, August, and September of 2018. It is shown that the process of interpolating PQPFs into the probability of exceedance values and then using them as an input to the hydrologic model produced forecasts that were able to capture the observed changes in the streamflow with a containing ratio of 100%. However, the low probability of exceedance values was associated with discharge values that were extreme, being ~34 times higher than average observed discharge. These high values are likely the result of the approach being too simplistic in that precipitation amounts for a specified exceedance value at every grid point, computed from the PQPFs and were then averaged and input into the Sacramento Soil Moisture Accounting model. Such

an approach assumes that all points in the basin would experience rainfall with potentially unusually heavy intensity and longevity. The error in the streamflow forecasts could be counteracted by calibration of the probabilistic derivate precipitation forecasts or by studying the typical distribution of precipitation within convective storms to adjust the rainfall inputs into the hydrologic model.

CHAPTER 1. GENERAL INTRODUCTION

Precipitation is one of the most critical factors for life on this planet as precipitation is the final step of the transportation of water vapor from the ocean onto the continents. It's what gives us the ability to grow crops, enjoy different types of recreation, and fulfill our daily sustenance needs. But it also can be dangerous, in the form of flash, in-land or coastal flooding, making it one of the costliness natural disasters that United States faces both in damages and in human lives (U.S. Billion-Dollar Weather and Climate Disasters, 2017). Additionally, it is also one of the hardest hydrometeorological events to forecasts even today, which is due to the difficulties that arise from both atmospheric models and streamflow models.

Atmospheric models have been continually improving in their ability to forecast weather events, whether it be in temporal or spatial resolution, or better techniques for modeling different atmospheric processes, such as heat and moisture transport or precipitation formation. There are still challenges in accurately predicting the location and magnitude of warm-season rainfall events (Ebert et al. 2003, Fritsch and Carbone 2004, Moser et al. 2015). These challenges are the result of many different factors including the small nature of the event, strength, and timing of dynamical lifting, the amount of energy present, and concentration of water vapor. All these factors must be considered by the model to generate an accurate precipitation forecast.

Additionally, hydrological models have also been improving in their ability to forecast streamflow changes. These advancements have occurred due to increases

in computational power allowing for higher resolution land models and smaller order streams to be used when modeling streamflow, and better routing techniques for water movement on the surface, sub-surface and in the channel (Singh 2018). Even with these advancements with streamflow forecasting, one of the biggest hurdles yet for the future is the combination of the two modeling systems.

Current streamflow forecasting practices are limited to using quantitative precipitation estimates (QPE) as problems associated with atmospheric models are thought to create too much inaccuracy in quantitative precipitation forecasts to be used as inputs (Georgakakos and Hudlow 1984, Lu et al., 2010). Additionally, the combination of quantitative precipitation forecasts (QPF) in hydrologic models, allows errors to continue to grow and increase the uncertainty in predicted discharges. This requirement of using QPE instead of QPFs limits the River Forecasting Centers (RFC) ability to issue warnings earlier as the RFCs needs to wait until the rainfall has already started.

The present study offers a possible method to generated streamflow forecasts before quantitative precipitation estimates could be measured. As inputs for the hydrological model, this study collected warm-season probabilistic quantitative precipitation forecasts (PQPF) generated from two different convective-allowing atmospheric models: High-Resolution Rapid Refresh Ensemble and the High-Resolution Ensemble Forecast version 2.0 model. By utilizing the ensemble's PQPF, the output of the model factors all the differences of the members' giving a more rounded solution. Additionally, these values can provide the probability of the event

occurring rather than a deterministic result providing more information to the streamflow forecaster.

CHAPTER 2: CREATE PROBABILISTIC STREAMFLOW FORECASTS USING HRRRE & HREF PROBABILISTIC QUANTITATIVE PRECIPITATION FORECASTS

by

Andrew R. Goenner, Kristie J. Franz, William A. Gallus, Jr.

Dept. of Geological and Atmospheric Sciences

Iowa State University, Ames, IA

and

Brett Roberts

Cooperative Institute for Mesoscale Meteorological Studies and

National Severe Storms Laboratory

University of Oklahoma, Norman, OK

A paper to be submitted to the Journal of Hydrometeorology

Corresponding Author: Andrew R. Goenner, 420 Birch St.,

Riverfalls, WI 54022, goena5759@gmail.com

2.1 Abstract

The present study examines how skillful probabilistic streamflow forecasts are when using convection-allowing ensemble models' probabilities of precipitation exceeding specified threshold accumulations as input. Both the High-Resolution Rapid Refresh Ensemble (HRRRE) and High-Resolution Ensemble Forecast version 2.0 (HREF) output were tested. A vital component of this work was the creation of

expected rainfall amounts at every grid point for seven different probability of exceedance values. The rainfall amounts for each of the probability of exceedance values were calculated using cubic interpolation from the probabilistic quantitative precipitation forecasts (PQPFs) generated from the HRRRE and the HREF models by use of a Gaussian smoothing technique applied by the HRRRE and HREF developers. The grid point precipitation amounts associated with the probability of exceedance values were then inputted into a hydrologic model for 11 different river basins across the upper Midwest for 109 cases during June, July, August, and September of 2018. It is shown that the process of interpolating PQPFs into the probability of exceedance values and then using them as an input to the hydrologic model produced forecasts that were able to capture the observed changes in the streamflow with a containing ratio of 100%. However, the low probability of exceedance values was associated with discharge values that were extreme, being ~34 times higher than average observed discharge. These high values are likely the result of the approach being too simplistic in that precipitation amounts for a specified exceedance value at every grid point, computed from the PQPFs and were then averaged and input into the Sacramento Soil Moisture Accounting model. Such an approach assumes that all points in the basin would experience rainfall with potentially unusually heavy intensity and longevity. The error in the streamflow forecasts could be counteracted by calibration of the probabilistic derivate precipitation forecasts or by studying the typical distribution of precipitation within convective storms to adjust the rainfall inputs into the hydrologic model.

2.2 Introduction

Floods are one of the most frequent natural disasters that occur in the United States, causing the loss of lives of 116 people and over $60 billion in damages in 2017 (NWS, 2018). In the Midwest in 2018, flash flooding caused damages to crop and property totaling over 100 million dollars and three deaths (Storms Events Database, 2018). Some of these losses could have been prevented if better practices were available to forecast flash flood events.

There are numerous inputs for modeling flooding events with one of the most important being atmospheric inputs, specifically rainfall. Unfortunately, precipitation forecasts have generally been thought to contain too much error to be used as inputs. Thus, flood forecasts are often not made until after precipitation has already fallen so that quantitative precipitation estimates (QPE) could be used as inputs instead of forecasts. Currently, the River Forecast Centers (RFCs) use QPEs generated from radar or rain gauge measurement (STAGE IV) as data for rainfall forcing for the hydrologic model (Nguyen et al. 2015, Krajewski et al. 2017). QPE, even though is represented as observed rainfall also has errors associated with it (Hou et al. 2014, Nelson et al. 2016). These errors are caused by issues such as inaccuracies in Z-R relationships (Wilson and Brandes, 1979), errors related to the angle or distance of the radar (Kitchen and Jackson, 1993), or the fact that rain gauges are only a point measurements and thus fail to capture the whole rainfall over an area (Villarini et al. 2008). All these factors impact the reliability of the rainfall forecasts. Another setback with using QPEs is since it is not generated until

after the rain has already begun, limiting the ability of forecasters to predict streamflow changes beforehand and shortens the amount of time that emergency managers have to inform and prepare safety personnel and the general public.

The accuracy of the atmospheric models' forecasts continues to improve both in timing and location, suggesting there may be ways to improve lead time for flood forecasts by using the atmospheric rainfall forecasts. Many studies have considered using Quantitative Precipitation Forecasts (QPF) as inputs to hydrologic models (Davolio et al. 2008, Cuo et al. 2011, Wu et al. 2014, and Seo et al. 2018). They have found that it is possible to use QPF as input for the hydrologic models' forecasts as they improved in spatial and temporal resolution. An additional study found that uncertainty increases as forecast times lengthen, limiting the range of time in which a QPF could be useful as an input for the hydrological model (Rabuffetti and Barbero, 2005). A final challenge is that warm season precipitation is typically associated with relatively small-scale thunderstorms that drop heavy rains that often leads to flooding (Ebert et al. 2003, Moser et al. 2015). The small scale and intense in nature of this rainfall aggravate errors in the atmospheric model forecasts that have serious consequences when QPF is used in hydrological models. The errors in timing and location for QPF are still usually large enough to have severe impacts on flood forecasts within small basins (Carlberg et al. 2018). These errors are caused by the atmospheric models' inabilities to accurately predict small-scale forcing (Golding 2000, Eckel and Mass 2005, Duda and Gallus 2013,) or complications arising from the use of parameterization of the boundary layer or microphysics schemes (Cintineo et al. 2014, Banks et al. 2016). These problems may

cause differences in the simulated convective from observed conditions (Avolio et al. 2017).

One possible way to mitigate these errors is to use the product of an ensemble of models since the ensemble provides a measure of uncertainty which is usually conveyed through probabilistic forecasts, and the mean, at least for many parameters, is typically more skillful than forecasts from any individual member (Leith 1974, Buizzza et al. 1999). In recent years there has been increasing focus on using Ensemble Prediction Systems (EPS) to assist with forecasting since EPSs can show the uncertainty and probability of an event occurring (Nielsen and Schumacher 2016, Clark 2017). Additionally, they provide more understanding to the user than single model systems (SMS) can (Ebert 2001, Leutbecher and Palmer 2007). This is a result of the use of the multiple models/members in the EPS, where each of the members has different initial conditions or dynamical and physical parameters, which can capture different nuances of the atmosphere and then be factored into the mean ensemble product; which an SMS is unable to do. There has been a study, Davolio et al. 2008, showing QPFs generated flood forecast using data from a multi-model were useful, but still had issues with modeling the timing and shape of the hydrograph of the river. While EPSs generally do benefit from higher rainfall location accuracy, one major problem that occurs when calculating the mean of a variable, such as QPF, during post-processing; it eliminates the extremes, possibly leading to missed high rainfall values that could occur over the basin that might generate a flood. A possible solution to this issue in eliminating the variability of the individual members would be the use of Probabilistic Quantitative

Precipitation Forecasts (PQPFs). PQPFs are generated based on how often each ensemble member exceeds a rainfall threshold. Post-processing is performed using statistical or neighborhood techniques to smooth the values, which still allows some of the high rainfall values to be present afterward in the probability values (Ebert 2009, and Gilleland et al. 2009, Schaffer et al. 2011). By looking at the likelihood of rainfall, we may be able to increase the value of the streamflow forecasts by making users aware of not only the most likely streamflow scenario but also the lower probability of extremes.

In the remainder of the paper, section 2.3 consists of a discussion of methodology, while section 2.4 will show the results from the streamflow forecasts, and section 2.5 presents the conclusions.

2.3 Data and Methods

2.3.1 Model and Probability values Setup

PQPF values were used as forcing for the hydrologic model and were obtained from two sources: The High-Resolution Rapid Refresh Ensemble (HRRRE) model of the Earth System Research Laboratory and the High-Resolution Ensemble Forecast version 2.0 (HREF) model of the National Centers for Environmental Prediction. The HRRRE ensemble consists of 9 members that run over a half CONUS domain with a resolution of 3 km. The members diverge by creating random variations in the zonal winds, temperature, and water vapor as part of the initial and boundary conditions of the individual members (Dowell et al. 2018). Each of the members used the same microphysics, Aerosol-Aware Thompson (Thompson et al.

2014), and boundary-layer scheme, Mellor–Yamada Nakanishi Niino (Nakanishi and Niino 2009). Random atmospheric perturbations were generated by using the first 36 members of the Global Data Assimilation System (GDAS) formatted to fit on the HRRRE domain. After formatting the GDAS data, random values of soil moisture were applied to the domain to generate variations in resulting forecasts (Trevor Alcott, NOAA, September 2018, personal communication). The 6-hour Accumulated Precipitation (APCP) was given at four different levels: 12.7, 25.4, 50.8, and 76.2 mm. Using these levels, it was determined how many of the members exceeded the thresholds to generate a probability value at each grid point. Afterwards, a Gaussian spatial smoother with a radius of 24 km from the grid point ran over the probability values to generate Probabilistic Quantitative Precipitation Forecasts.

The HREF is a time-lagged ensemble built using four models, the first two being variations of the Weather Research and Forecasting (WRF) model: High-Resolution Window (HRW) National Severe Storms Laboratory model (NSSL) and HRW Advanced Research WRF model (ARW). Additionally, there are also two variations of the NAM model: the NOAA Environmental Modeling System's Nonhydrostatic Multiscale Model on the B-grid (NMMB) and the North American Mesoscale (NAM) Nested forecasting model (Roberts et al. 2018). Model forecasts are generated twice a day on a full CONUS domain from the models above at 00z and 12z. In addition to the four model runs mentioned above, the HREF uses four time-lagged members from the previous models that occurred 12 hours prior. In total, there are eight members in the HREF: four current time and four time-lagged members. Between the four different models, there were two types of boundary

layer schemes, with both NAMs and the HRW NSSL using the Mellor-Yamada Janjic scheme (Janjic 1994), while the HRW-ARW used the Yonsei University Scheme (Hong et al. 2006). Additionally, the two WRF models used two microphysics, WRF Single-moment 6-class scheme (Hong and Lim 2006) while the NAM models used Ferrier-Aligo scheme (Aligo et al. 2014). Table 1 shows both HRRRE and HREF model setup. The rainfall thresholds for which the HREF generates probability values were 6.4, 12.7, 25.4, and 50.8 mm. HREF creates the PQPF values is similar to the HRRRE; the probability of each grid point exceeding that threshold was calculated, and then a Gaussian smoother was ran over the resulting field, where the probability values are counted within 40 km of an individual grid point (Roberts et al. 2018). The only difference between the HRRRE and HREF is that the radius values used for HRRRE were 24 km or 8 grid points compared to 40 km or 13 grid points for the HREF.

To derive the rainfall amounts at the desired probability levels, a cubic interpolation was conducted over the input levels to extract precipitation amounts at specific probability values of 5, 10, 25, 50, 75, 90, and 95%. These values were selected as they would match the Probability of Exceedance values used at the River Forecast Centers for discharge. To complete the cubic interpolation, the function requires at least three unique points. HRRRE and HREF generate probability forecasts at four different rainfall amounts, (12.7, 25.4, 50.8 and 76.2 mm, and 6.4, 12.7, 25.4, and 50.8 mm, respectively). Both HRRRE and HREF regularly had predictions of zero probability of occurrence at rainfall amounts equal to or greater than 25.4 mm, thereby limiting the number of unique data points needed to

complete the interpolation. As a result, it was necessary to expand the number of unique probability values to be able to achieve interpolation. Additionally, it was desired to remove the need to use extrapolation on the edges of the PQPF values, as extrapolation is known to have the possibility of giving answers that are not realistic. Figure 1a show an example of the interpolation process with a set of HRRRE values.

The first additional point added was a precipitation value of zero. At the rainfall amount equal to zero, it is guaranteed that at least zero rainfall will occur whether it is a day with a high probability of heavy rainfall or one with clear weather, so a value of 100 percent was assigned to this data point. This addition brought the number of data points up to 5 and closed the range of values for the interpolation on the low end. A second additional point was added to the probability data by looking in the opposite direction at the maximum rainfall in the domain shown in Figure 2. The process to find this point was done by first using the weighted APCP values (being the highest member rainfall value at every grid point). Afterwards, the highest APCP grid point inside the domain was selected, and a hundredth of an inch of rainfall was added to this value, making it impossible to find this value inside the domain; thereby, a probability of zero was could be assigned to this rainfall amount. This zero value could be used to close the interpolation of the probability rainfall amounts on the upper end of the dataset. Since there were now six unique data points and the full range of probability values enclosed, the cubic interpolation could be used to derive our probability of precipitation values, as shown in Figure 1a.

Three cubic interpolation programs could have been used in the study: Piecewise Cubic Hermite Interpolating Polynomial (PCHIP), Modified Akima Cubic Hermite Interpolation (Makima), and not-a-knot spline (SPLINE). Azizan et al. 2018 found when comparing the PCHIP and SPLINE on interpolated rainfall values; the SPLINE could produce negative rainfall values compared to the PCHIP which stayed in the positive. Additionally, the computational power to interpolate with PCHIP was stated to need lesser computational resources than the other two forms of interpolation systems. So, it was decided to use the PCHIP interpolation program to predict the probability values similar to those seen in Figure 1b.

There were still occurrences when the only unique data points were 100 percent at zero rainfall and 0 percent at 6.4 and 12.7 mm in the atmospheric models. In these occurrences, the values of 2.54, 2.29, 1.78, 1.27, 0.76, 0.51, and 0.25 mm were assigned to the 95, 90, 75, 50, 25, 10, and 5 percentiles, respectively. It was expected that these occurrences would play little to no role in the later calculations of the basin averaged precipitation, as the scope of the project was focused on peak flows generated by heavy precipitation events. After generating the precipitation amounts at the different probability values, the precipitation amounts were extracted from the domain so that only the grid points located inside or contacting the edges of the basins were selected. These selected grid points were then averaged to determine a basin averaged precipitation value to be fed into the hydrologic model, Figure 3.

2.3.2 Study Basins case selection

The study area was the Upper Midwest including the states of Illinois, Iowa, Minnesota, and Wisconsin, Figure 2. This region is characterized by forested hills and lakes in the northern parts of Minnesota and Wisconsin to plains and farmlands in Illinois and Iowa. On many of the river basins, minor urban development has occurred including the cities of Ames, Iowa, and Waukesha and Kenosha, Wisconsin. To be consistent with previous studies done by Dziubanski and Franz 2016 and Carlberg et al. 2018, eleven different basins were selected from the Upper Midwest region, Table 2. Also, the gauge locations were chosen to be the uppermost gauge of the rivers to simplify calculations, as it would limit the need to account for inflow of water from upstream gauge locations when modeling the watersheds. When performing statistical analysis, the HRRRE and HREF forecasts were compared to those made by the North Central River Forecasting Center (NCRFC), the streamflow events were limited to just the gauge locations of the NCRFC forecast.

Different rainfall/discharges events were selected based upon a substantial increase in peak discharge values when compared to the base flow values present before the start of rainfall. Additionally, the observed max discharge needed to reach a discharge value higher than 75 percent of occurrence based on climatology data during the warm months (June, July, August, and September) generated from United States Geological Survey (USGS) discharge measurements, reach action stage, or have flooding conditions that were present during the model forecasts period.

2.3.3 Hydrologic Model

To be as similar to the NCRFC's flood predictions as possible, the present study used the spatially lumped SACramento Soil Moisture Accounting model (SAC-SMA) (Burnash et al. 1973). The model was developed in the 1960s at the River Forecasting office located in Sacramento, California. Currently is being used as the operational hydrologic forecasting models of the RFCs. The primary purpose of the model is to determine the amount of runoff that will occur in a catchment (watershed) as a measure of expected discharge. The principle that the SAC-SMA model uses to calculate streamflow changes is the water balance equation:

Runoff (Discharge) = Rainfall – Evapotranspiration – Change in Soil Moisture. (1)

Each of the four parts of the water balance equation has been quantified into different parameters or states describing how water storage changes. Precipitation inputs being the main forcing for the model. These states are based on basin characteristics, current soil moisture, and unit hydrograph. Some of the states are stationary and set for a catchment (soil type, impervious runoff amount, unit hydrograph at the gage location, and monthly evapotranspiration rates), while others change dynamically based on the current conditions (soil moisture conditions, upper zone water values, free water, etc.). Both types of states are then used to calculate a given change in the water balance equation in a catchment to describe the discharge at a given forecast points.

For the project, a warm start was used on the SAC-SMA model employing initial condition values from the NCRFC to mimic their forecasts. In initial testing, it

was found that the SAC-SMA discharge simulations were significantly below observed values for the first few timesteps; after 4-8 timesteps, the stream simulated discharge would reach near-observed discharges. To minimize errors caused by the presence of low discharge at the beginning of the forecast, the study initiated the model runs with a five-day spin up period at the start of the streamflow forecast. When running the averaged probability-based precipitation values through the SAC-SMA, an additional run was completed using STAGE IV, measured precipitation data, from the NCRFC.

2.3.4 Statistical Analysis Approach

Analysis of the discharge data for this project was completed using bias, percentage bias, containment ratio, percentage frequency, and rank probability scores. The relative percentage bias was calculated by taking the difference of an individual observed peak discharges (O_{peak}) and an individual peak discharge of the forecast (F_{peak}) for each probability, event and basin, which was then divided by (O_{peak}), (Equation 2). Afterwards, these individual percentage biases were averaged across basins and events per the different forecasted probability levels.

$$\text{Relative percentage bias} = \frac{O_{peak} - F_{peak}}{O_{peak}} \quad (2)$$

Containment ratio was used to examine whether forecasts were able to contain the observed peak discharge values within the different probability of exceedance forecasts. If the result of the ratio is one or close to, the ratio shows that forecasts were able to exceed the observed peaks while scores near 0 indicate that

forecasts regularly underpredicted the peak observed discharge and were unable to contain them.

Ranked Probability Scores (RPS) have been used in the past to examine meteorological and hydrologic forecasts (Franz et al. 2003, Wilks 2011, and Carlberg et al. 2018). The benefits that RPS gives over other types of statistical analysis is that it can evaluate multiple categories of probability values (Wilks 2011). RPS is similar to Brier Score in that they both calculate the mean squared errors, but a Brier score is limited to one yes-no answer, while RPS tests over multiple categories at one time. To calculate the RPS, four different exceedance categories of discharge were selected: flood event, where the peak discharge reached or surpassed the minor flood discharge; action stage when the discharge was higher than action stage discharge but lower than minor flood discharge; 50% of action stage, the level of discharge was greater than half of the action but less than action stage; and nonevent, where the discharge was larger than zero and less than 50% action stage. These exceedance levels were chosen to be similar to the study done by Carlberg et al. 2018. To compare the RPSs from HRRRE and HREF rainfall forced discharges, the study requested the probabilistic discharges forecasts generated by NCRFC.

NCRFC generates their probability of discharge prediction (at values of 5, 50, and 95% chance of exceedance values) based off an ensemble of 46 members produced by the Weather Prediction Center (WPC). This ensemble includes members from SREF (both ARW and NMMB), GEFS, ECMWF ensembles, GFS, NAM,

WRF Hi-res ARW, WRF Hi-Res NMMB, ECMWF models, and WPC's own generated model forecast. Most of the members for the HREF are included with the exception of the time-lagged members. WPC creates these forecasts at 9z and 21z daily. NCRFC generates the probability forecasts daily at 18z forecasting out to 168 hours. Thus, NCRFC use the 9z probability forecasts from WPC to make the probability forecasts. In total 79 forecasts were compared across the seven shared basins, as currently, the NCRFC is unable to generate probability discharge forecast for all catchments in the Upper Mississippi region.

Although an apples-to-apples comparison of forecasts was not possible because of the different times at which the forecasts were made, to have the most appropriate comparison possible, every effort was used in selecting the NCRFC probability forecasts closest to the HRRRE and HREF run times. For example, if a 00z ensemble was to be used, the 18z NCRFC probability forecast generated six hours beforehand was selected. When using the 12z ensemble runs, an examination of archived radar imagery from University Corporation for Atmospheric Research (UCAR) (UCAR Image Archive, 2018) was examined to see if any significant rainfall occurred between the 12z and 18z period. If there was no rainfall in the basin during this period, the NCRFC probability forecasts generated at 18z, six hours after the HRRRE and HREF initialization, was selected, this occurred in 41 events. If there was rainfall in the basin before 18z, the previous 18z forecast, 18 hours prior, was used, this happened 15 times over the course of the study.

It is acknowledged that this approach often resulted in selecting NCRFC forecasts issued prior to the HRRRE and HREF forecasts, causing NCRFC forecasts to have longer lead times, which tends to worsen forecast performance. However, the NCRFC forecasts are based on WPC forecasts that use a much larger ensemble of 46 members compared to HRRRE's 9 and HREF's 8; additionally, the WPC forecasts benefit from the expertise of the forecasters there.

The process that the RPS is calculated is by the following equation (3), where n is the number of categories, F is the forecasted probability value at an exceedance discharge category, and O is either the value of 0 or 1 depending on whether the observed peak discharge reached occurred at that discharge categories. Using the different discharge categories, the probability of each categorical level was calculated using 5, 50, and 95% discharge amounts. To find each probability of the event categories, a linear interpolation was used to interpolate between the categorical levels, and the resulting probability values were used as input for F in equation (3).

$$\text{RPS} = \frac{1}{(n-1)} \sum_{1}^{n} (F - O)^2 \quad (3)$$

RPS values range from 0 to 1, with 0 being the best as the model forecasts were able to assign the correct probability values to the corresponding exceedance categories of discharge.

2.4 Results

2.4.1 General Results from the Eleven Basins

A total of 109 events were examined across the 11 basins (79 events across 7 basins when compared to NCRFC), with 33 (26) of the events exceeding action stage, Table 3. A general pattern in errors in discharge prediction was observed to occur between the two model-derived rainfalls amounts at multiply probabilities of exceedance values. Between the two ensembles, the HRRRE had an average bias of 818 cms (3,330%) while HREF had an average bias of 810 cms (3,470%) for forecasts of 5% exceedance, Table 4. At 95% exceedance, and HRRRE had average errors of -28.1 cms (-43.5%) to HREF's -17.3 cms (-20.7%). The probability of exceedance value that was most similar on average to the observed discharge was the 90% exceedance for the HREF and 75% exceedance for the HRRRE. As the value of the probability of exceedance approached 50%, HREF's error in discharges (189 cms) tended to be twice as large as those of the HRRRE (95.7 cms). The most probable reason for this difference in probability values, was HREF having higher values of accumulated basin averaged rainfall across the different probability thresholds than HRRRE, thereby producing higher discharge amounts for each discharge probability, Table 4. As seen in Table 4, between HRRRE and HREF similar biases were observed at the high and low exceedances; this is the result of both the HRRRE and HREF probability rainfall inputs being similar, Table 4. When examining the high probability of exceedance, a large percentage of the rainfall would be transferred into ground limiting the amount of water to output as discharge. While

at the low exceedance values the model mostly suffered from trying to output 200+ mm of rainfall resulting, the SAC-SMA model generating similar discharges. When observing the 25%-75% probability values, it is seen that where the atmospheric models had differences in their rainfall inputs, subsequently producing a similar variances in their discharge outputs.

The ranked probability scores showed that HRRRE tended to have better scores on average compared to HREF, Figure 4. The average RPS for HREF was 0.36 (standard deviation, 0.09) compared to HRRRE's average RPS of 0.29 (standard deviation, 0.06) (Table 5). This difference in RPS values is most likely due to HREF producing higher rainfall amounts compared to HRRRE thereby shifting probability values and resulting in more error when computing the RPS with the four discharge categories mentioned in the methodology. When comparing the HRRRE and HREF to the NCRFC, it was seen that both ensembles had better RPSs than the NCRFC. The average RPS for the NCRFC was 0.59 with a standard deviation of 0.07 (Table 5). This increase was due to the NCRFC's discharges tendency to underpredict the observed discharge at multiple probability values. This inclination to under-predict, in turn, lowers the probability values assigned to discharge categories during the RPS calculation and overall increased the RPSs.

RPSs between HRRRE and HREF averaged in the 0.3 to 0.4 with the exception the Root River at Pilot Mound, MN (RPMM5) which had values of 0.49 and 0.61 for HRRRE and HREF. It is believed that this increase in error was due to the fact the RPMM5 was the largest basin studied, allowing for more water to accumulate as a

result of the rainfall distribution. During the basin averaging process the probability of exceedance rainfalls values would be distributed across all grid points within the basin and all timesteps of the forecast. When the averaging was completed and inputted into the SAC-SMA model, the model then had to account for and calculate the discharge at the gauge location for this possibly unreasonably large rainfall which in turn generated unrealistically large discharges. This process of average all the 5% exceedance rainfall amounts inside the basin would allow the discharge values to grow to a substantial amount even though the stream never reached action or flood stage during the study, Table 3. Which in turn would place a high probability of exceedance values in the flood category, raising the RPS for the basin.

When comparing the ability of the three forecasts to capture the relative frequency of the observed peak discharge, it was observed that HRRRE was the most accurate at forecasting the 95% exceedance value, followed by just behind by the NCRFC, Figure 5. At the higher probabilities, the NCRFC forecasts tended to heavily under-predict the frequency of peak discharge amounts at 50% and 5% exceedance values compared to the HREF and HRRRE. HRRRE did a better job overall at predicting the frequency of occurrences for the different probability values, landing just under the intended mark at 95% and 90% and then falling to a larger degree at 75% through 5%. The tendency for HREF to overpredict the rainfall at all probability values and causing the observed discharge to exceed the forecasted discharge, forcing the frequency scores into the high 70s at the 95% and continuing onto zero at 25%. Also, when looking at the containment ratio for the probability forecasts, both the HRRRE and HREF were able to capture the observed discharge

100% of the time, compared to the NCRFC forecasts which did so only 60% of the time.

Now that we have seen how the broad results generated from the probability of exceedance forecasts preformed in predicting streamflow changes, two example case studies were selected to give a more in-depth discussion of the forecasting ability. The first case study, June 14th along the Skunk and Squaw Creek at Ames IA, was selected as it was a rapid increase in streamflow, caused by heavy convective rainfall. The second case study, September 30 - October 1 along Volga River at Littleport IA, was a rapid and large increase in streamflow, associated with moderate rainfall in exceptionally wet condition.

2.4.2 Case Study June 14th, Skunk River and Squaw Creek

One of the events showing that the technique that was developed could be useful in flash flood prediction was a flash flood event that occurred in the city of Ames, Iowa on June 14th, 2018. Ames is located at the junction of the Squaw Creek with the Skunk River. The Skunk and Squaw Creek watersheds are similar as both basins are situated inside the Des Moines Lobe and the primary land cover type is cropland of corn or soybeans with minor urban development at the rivers' convergence and the surrounding area. The two USGS gauges [05470000/AMEI4 (Skunk above confluence)), 05470500/AMWI4 (Squaw above confluence)] used in the study are located around the city. Both streams were at or below the median discharge on June 13th (USGS Skunk and USGS Squaw) at ~4.2 cms.

In the early morning hours (7 UTC) of June 14th, a line of multicellular convection in connection with a cold front moved over the watersheds of the Skunk and Squaw and produced heavy rain over the basins for the next twelve hours. According to Iowa Mesonet dataset, the system deposited 107 mm of rainfall with a peak rainfall rate of 40.9 mm an hour. Other gauges monitored by Community Collaborative Rain and Hail Snow Network (COCORAHS) volunteers in the area measured amounts as large as 178 mm. To model this event, the June 14th 00Z model runs were selected from both the HRRRE and HREF. Stage IV data show that heaviest total rainfall occurred over the Squaw Creek basin, with the majority of the 6-hour accumulation occurring between forecast hours 12 – 18, June 14th 12-18z. As a result of this short heavy rainfall, flash flooding occurred in the city of Ames causing streets to be closed and people needed to be rescued. As shown in Figure 6, both the HRRRE and HREF PQPF values were able to predict that rainfall was going to occur in the region of the Skunk and Squaw basins. With HRRRE's probability forecasts being adjusted slightly to the north of observed STAGE IV precipitation and HREF's being centered on the basins. Both HRRRE and HREF were able to produce the same shape and direction of the observed rainfall.

For both the HRRRE and HREF, 5% through 90% exceedance probability fields were able to produce some increase in discharge, with HREF's 95% exceedance values also producing an increase in the discharge for both basins. The observed peak discharge for the Skunk River was 89.2 cms (action stage is 122 cms) while Squaw Creek's peak discharge was 120 cms (action stage being 108 cms). In both sets of hydrographs (Figures 7a and 7b), they produced a similar flood wave

both in shape and timing to the observed. When comparing the HRRRE and HREF for the Skunk River, the probability value that was the most similar to the observed discharge for both models was the 50% exceedance probability falling just below the observed discharge with both models under-predicting a few cms, (Table 6). The most significant difference between the two models was that the HRRRE's 95% exceedance discharge didn't increase in discharge during the time period of the forecasts compared to the HREF discharge increase of 6 cms. The STAGE IV, precipitation for the Skunk River, over predicted the observed stream change having peak discharge of 86.0 cms above the observed. The NCRFC probability forecast for the Skunk River, were tiny with NCPFC 5% exceedance predicting a discharge of 6.40 cms almost 14 times smaller than observed (Table 7).

Over the same forecasting period, the Squaw Creek basin received a higher amount of rainfall, which caused the river to exceed action stage at the gauge location in Ames. Here again, both hydrographs showed that the generated flood peaks from the ensemble model interpolated rainfall values were able to model flood wave both in shape and size, Figure 8. Again, the HRRRE's 95% exceedance values didn't produce an increase in discharge, due to HRRRE having smaller rainfall amounts than HREF. For HRRRE the 25% exceedance peak was the probability values that was the closest to the observed peak while for HREF the 50% was the most similar to the observed discharge (Table 6). When looking that the NCRFC forecast here, the predicted forecast for 5% exceedance was again small, only forecasting a discharge of 10.3 cms (Table 7). These results show that it is possible to use the ensembles' PQPF to generate stream peak changes associated with flash

flood before rainfall had started thereby allowing local emergency managers the possibility of close areas prone to flooding.

2.4.3 Case Study September 30th -October 2nd, the Volga River

The second case occurred along the Volga River basin at the USGS location (05412400) at Littleport, Iowa September 30th through October 2nd. The area of northeast Iowa, located next to the Mississippi River, is more rugged then the plains location of Ames, Iowa as the region was not impacted by the last glacier movement, leaving the area to be cut by numerous different rivers. Each is separated by ranges of hills and cliffs which limits drainage to occur in a northwest to the southeast direction towards the Mississippi. As a result of the massive and numerous rainfall events that were happening in and around the basin during September, the Volga River in Littleport was running at ~28.3 cms compared to the average discharge of 2.83 cms at the end of September (USGS Volga 2018).

Throughout September 30th to October 2nd, a stationary front developed and stretched from the Colorado Rockies through the state of Iowa from the southwest to northeast, and onto the southern Great Lakes region. To the south, a high-pressure system centered over the southeast allowed moisture transport from the Gulf of Mexico to the Volga region. Moreover, a strong cold airmass was located over Minnesota allowing the warm, moist air to rise and initiate convection along the stationary front. Rainfall occurred for an extended period with the rain starting late on September 30th and continuing until the early hours of October 2nd. The Stage IV rainfall total for the USGS gauge was 114 mm with the peak single hour

accumulation being 21.0 mm. For this event, the HRRRE, and HREF's September 30th 12z model run was selected as it was able to capture the beginning and most substantial part of the rainfall while only missing parts late on October 2nd, which played a minimal role in the flood peak creation. The heaviest 6-hour accumulation occurred during forecast hour 24-30, October 1st 12z-18z. As seen in Figure 9, both HRRRE and HREF had the highest probability for heavy rainfall shifted to the east, into southern Wisconsin and northern Illinois. HREF was shown to be wetter and more accurate with the higher probabilities by having a likelihood of 50.8 mm occurring in the same area of observed 50.8 mm contour, in Figure 9g.

This event was noteworthy as it was the largest, and one of the quickest, increase of discharge between base flow and peak discharge with a shift of 255 cms in 18 hours. Both the HRRRE and HREF's probability forecasts were able to exceed the flood peak with their 5% chance of exceedance. Table 8 shows that the probability values that had the most similar observed discharge for both HRRRE and HREF were 25% exceedance. Both sets of atmospheric models PQPFs producing similar hydrographs with the only major difference between them being 95% exceedance discharge for HRRRE didn't produce an increase in discharge, Figure 10. Additionally, the 5% and 10% of exceedance discharge were still much higher when compared to the observed discharge, overpredicting the observed by hundreds of cms. These results show that there still is work needed to be done to lower the discharge amounts associated with the low exceedance values. This basin is not one of the forecasting locations for the NCRFC so it is unknown if the NCRFC probability forecasts would have been able to capture the flood event. This event shows that the

process that was developed can work even in some lowest probability events, as the method was able to capture and exceed the observed discharge and give the likelihood of the observed streamflow changes that did occur.

2.5 Conclusion

109 events, across 11 basins throughout the Upper Midwest, were used to examine a technique to forecast streamflow changes using Probabilistic Quantitative Precipitation Forecasts taken from two different convective-allowing atmospheric ensemble models, HRRRE and HREF. These PQPF values were then processed through an interpolation system to create a range of rainfall probabilities that were used as input to the spatially lumped Sacramento Soil Moisture Accounting model. A variety of different statistical techniques were used to analyze forecasts of peak streamflow and compared to current probability forecasts generated by the NCRFC.

The results show that there were differences in the changes of streamflow forecasts from the HRRRE, HREF, and NCRFC, with HRRRE, tending to have the most significant ability to make accurate predictions. This was shown in the RPSs and the relative frequency of forecast to exceed the observations, where HRRRE had the lowest average RPS of 0.32 and regularly was in better agreement with the expected relative frequencies at multiple probability values. It is believed that this is the result of the HRRRE's ability to have the probability basin average precipitation amounts that were lower and more similar to the observed rainfall. While HREF was also able to capture all observed changes in discharge, its tendency to generate higher discharge amounts at all probability values compared to HRRRE, in turn,

raised its RPSs and lowered relative frequency of forecast scores. Finally, NCRFC forecast was limited in the amount of the discharge produced at the different probabilities values generated thereby causing it to underpredict observed discharge, increasing the RPSs and raising the relative frequency scores.

Repeated issues were seen in both the HRRREs' and HREFs' low probability of exceedance discharges surpassing realistic discharge amounts for both models. This problem could be reduced in the future by calibrating the probability of exceedance values, to decrease rainfall amounts thereby limiting magnitude of predicted discharge. Another solution would be analyzing the distribution of the precipitation when extracting the precipitation values from the basins and arranging the probability exceedance values based on the spatial precipitation distribution before completing the basin averaging. This finding is echoed by Cuo et al. (2011) where they state that postprocessing may need to be completed on rainfall forecast data from both deterministic and ensemble forecasts to make them useful to predict streamflow changes. When looking at the ability of the technique to forecast streamflow changes, the two different case studies show that whether the event be a flash-flood or low probability of occurrence flood. The process would have been able to give emergency personal and the public more information about the possibility streamflow changes then current flood forecasting practices.

2.6 Acknowledgements

This research was funded by the National Oceanic and Atmospheric Administration C-STAR Award, NA17NWS468000. The author would like to thank

NCRFC, especially Brain Connelly, for providing the initial conditions data for the SAC-SMA model and the probability forecasts from the NCRFC.

2.7 Tables

Table 1. Member configurations for the HREF and HRRRE model during the period of research, including differences in initial conditions, lateral boundary condition, microphysics and planetary boundary layer scheme, grid spacing and number of vertical levels.

Member	ICs/LBCs	Microphysics	PBL	Grid spacing	Vert. levels
HRW NSSL (-12h)	NAM/NAM -6h	WSM6	MYJ	3.2 km	40
HRW ARW (-12h)	RAP/GFS -6h	WSM6	YSU	3.2 km	50
HRW NMMB (-12h)	RAP/GFS -6h	Ferrier-Aligo	MYJ	3.2 km	50
NAM CONUS Nest (-12)	NAM/NAM	Ferrier-Aligo	MYJ	3 km	60
HRRRE 9-members	GDAS (with random permutations added)	Thompson aerosol-aware	MYNN	3 km	51

Table 2. Location of the gauges of study watersheds, their area, USGS Gauge Number, North Central River Forecast Center identification name, and the number of selected events per a basin during the study. Bold signifies the gauges that NCRFC perform probability forecasts.

Basin	Area (USGS)	USGS Gauge Station	River Forecasting Center IDs	Number of events
Kickapoo at Ontario (WI)	303 km^2	05407470	ONTW3	10
Fox River at Waukesha (WI)	326 km^2	05543830	WKEW3	11
Turkey River at Spillville (IA)	458 km^2	05411600	SPLI4	11
Squaw Creek at Ames (IA)	528 km^2	05470500	AMEI4	14
Pecatonica East at Blanchardville (WI)	572 km^2	05433000	BCHW3	14
Pecatonica West at Darlington (WI)	707 km^2	05432500	DARW3	8
Des Plains at Russell (IL)	785 km^2	05527800	RUSI2	6
South Skunk River at Ames (IA)	816 km^2	05470000	AMWI4	14
Wapsipinicon River at Tripoli (IA)	896 km^2	05420680	TLPI4	7
Volga River at Littleport (IA)	901 km^2	05412400	VLPI4	6
Root River at Pilot Mound (MN)	1463 km^2	05383950	PRMM5	8

Table 3. Dates and start times of the model runs for each of the basins. Along with peak observed discharge where NEN represents, peak discharge fell between 0 and 50% action stage; >50 indicates, peak discharge between 50% action and action stage; ACT a peak discharge between action stage and minor flood; and FLD a, peak discharge at or greater than minor flood stage. Bold signifies basins that NCRFC generates probability forecasts.

Date and start time of the model run	ONTW3	**WKEW3**	**SPLI4**	**AMWI4**	**BCHW3**	**DARW3**	RUSI2	**AMEI4**	**TPLI4**	VLPI4	RPMM5
June 14th 00z				FLD				>50			
June 15th 12z		NEN									
June 20th 12z								>50		>50	
June 21st 00z			>50		>50	NEN	FLD		>50		NEN
June 25th 12z	NEN	>50	>50	NEN	>50	NEN	FLD	NEN			NEN
June 29th 12z	NEN										
June 30th 12z			>50	>50	>50			>50	>50	NEN	NEN
July 12th 12z	NEN										NEN
July 13th 00z					NEN						
July 19th 00z	NEN				NEN	NEN					
July 19th 12z		NEN			NEN						
August 1st 12z	NEN				NEN						
August 6th 12z				NEN							
August 7th 12z			NEN								
August 14th 12z				NEN	NEN			NEN			

Table 3. Continued

August 19th 12z	NEN	>50		NEN	ACT	NEN		NEN			
August 26th 12z		>50									
August 27th 12z		ACT	>50								
August 28th 12z				NEN	NEN		NEN	NEN	>50		
August 31st 00z			>50								
August 31st 12z				NEN				NEN	FLD		
Sept. 2nd 12z		ACT	ACT		>50	>50	>50			FLD	
Sept. 3rd 12z										FLD	
Sept. 4th 00z	FLD										
Sept. 4th 12z		FLD	FLD	>50			>50	ACT	FLD		NEN
Sept. 5th 00z				>50	ACT	ACT		ACT			
Sept. 18th 12z				NEN				NEN		FLD	
Sept. 19th 00z		NEN	FLD	NEN	ACT	ACT		NEN			
Sept. 19th 12z	>50						NEN		FLD		NEN
Sept. 20th 00z				NEN				NEN			
Sept. 21st 00z	>50										NEN
Sept. 30th 12z	NEN	ACT	NEN	>50	FLD	FLD		>50	NEN	FLD	
October 1st 12z		ACT									
October 2nd 00z											NEN

Table 4. Average bias in $m^3 s^{-1}$, average relative bias (in percent), and average rainfall in mm for all 11 basins and 109 events during the warm season of 2018 for a range of probability of exceedance values.

Bias	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**
HRRRE	818	566	283	95.7	9.46	-20.9	-28.1
HREF	810	666	422	191	53.5	-2.83	-17.3
Percentage	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**
HRRRE	3,330%	2,150%	1,000%	321%	47.6%	-26.5%	-43.5%
HREF	3,470%	2,720%	1,500%	619%	175%	15.8%	-20.7%
Rainfall	**5%**	**10%**	**25%**	**50%**	**75%**	**90%**	**95%**
HRRRE	226	176	119	73.9	40.1	21.8	14.0
HREF	249	215	159	103	61.2	34.8	23.1

Table 5. Average and standard deviation for HRRRE, HREF and NCRFC's RPSs for the 7 shared (all 11) basins.

	Average RPS	**Standard Deviation RPS**
HRRRE	0.29 (0.28)	0.06 (0.10)
HREF	0.36 (0.34)	0.09 (0.12)
NCRFC	0.59	0.07

Table 6. Bias for Skunk and Squaw Creek on June 14th at different probability of exceedance values in $m^3 s^{-1}$.

Skunk	**95%**	**90%**	**75%**	**50%**	**25%**	**10%**	**5%**	**STAGE IV**
HRRRE	-185	-185	-171	-137	94.3	371	615	86.0
HREF	-179	-167	-150	-64.6	220	411	507	
Squaw	**95%**	**90%**	**75%**	**50%**	**25%**	**10%**	**5%**	
HRRRE	-86.7	-85.2	-66.0	-19.5	157	402	620	20.2
HREF	-86.4	-77.9	-54.7	-9.60	163	430	561	

Table 7. NCRFC probability forecast for the Skunk and Squaw Creek on June 14th at different probability of exceedance values in $m^3 s^{-1}$

	95%	**50%**	**5%**
Skunk	5.70	5.70	6.40
Squaw	5.32	5.32	10.3

Table 8. Bias for the Volga River on September 30th and October 1st at the different probability of exceedance values in $m^3 s^{-1}$.

	95%	**90%**	**75%**	**50%**	**25%**	**10%**	**5%**	**STAGE IV**
HRRRE	-270	-267	-240	-90.6	219	535	700	-37.3
HREF	-230	-205	-31.2	134	356	645	969	

2.8 Figures

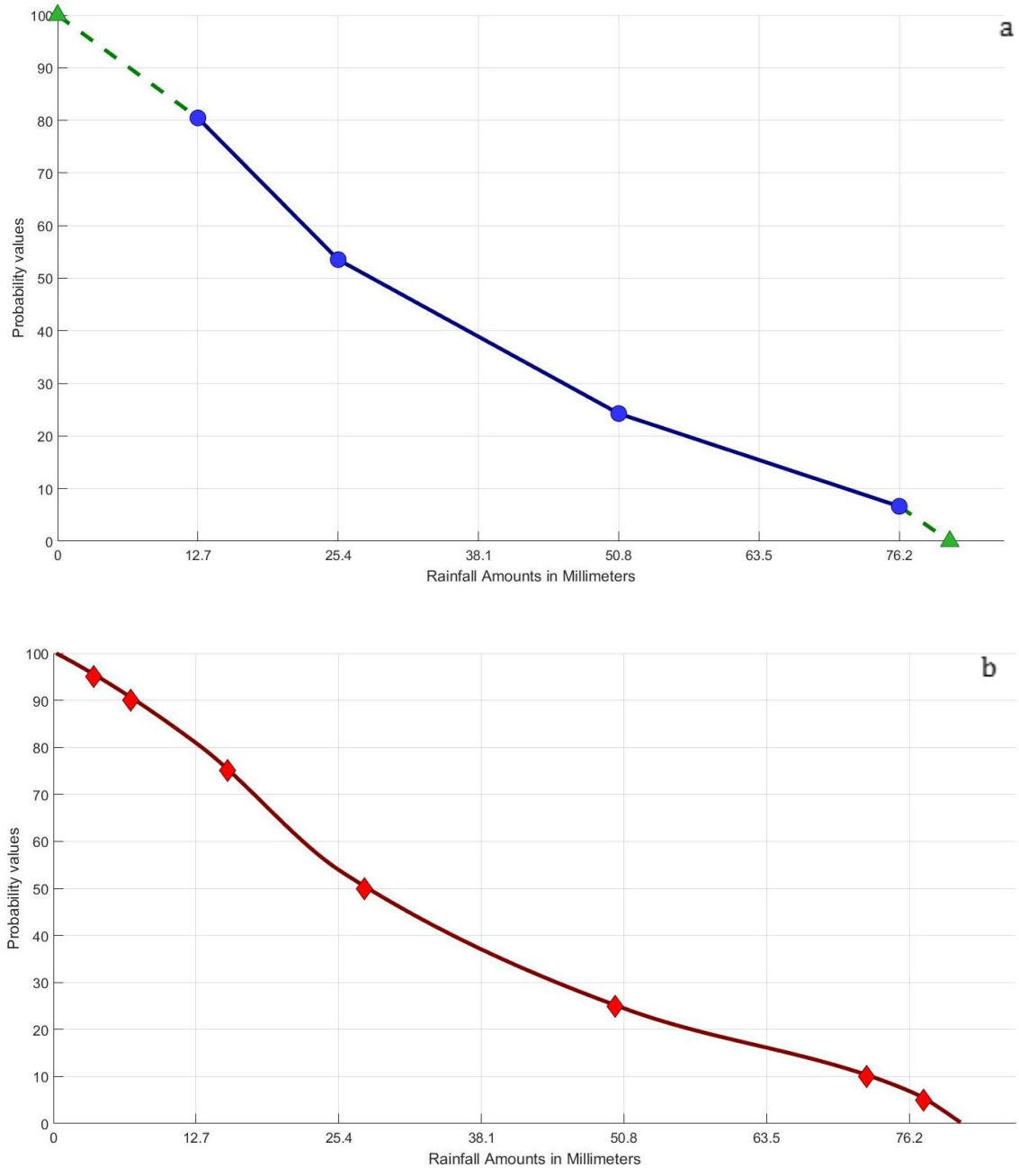

Figure 1: Example of how the PQPF data is interpolated and transformed into the rainfall amounts used as inputs into the hydrology model. The Figure 1a shows the PQPF values from the HRRRE (Blue Circles) with the additional boundary values (Green Triangles) from zero rainfall and the absolute highest rainfall amount inside the research domain. The Figure 1b shows rainfall amounts after running the PCHIP interpolation program at selected probability values (Red Diamonds) used in the study.

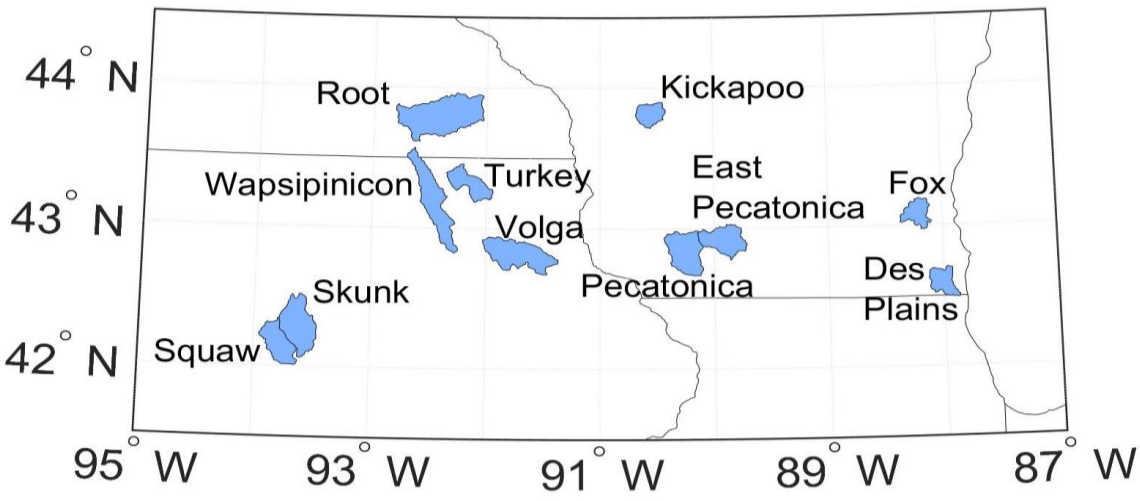

Figure 2: Map of the studied watersheds across the Upper Midwest.

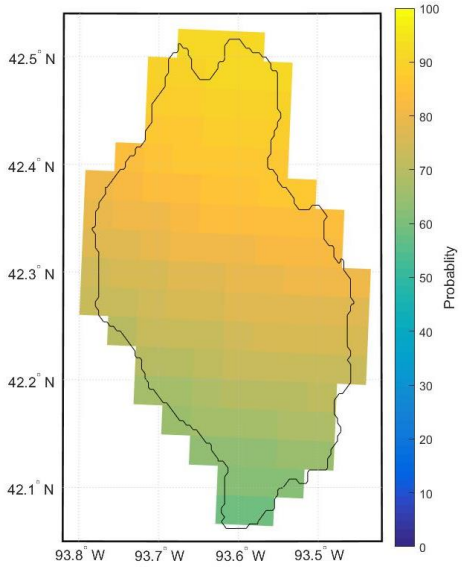

Figure 3: An example of clipping the HRRRE's probability values for exceeding 12.7 mm of rain over the Skunk River on June 14th from 12-18z.

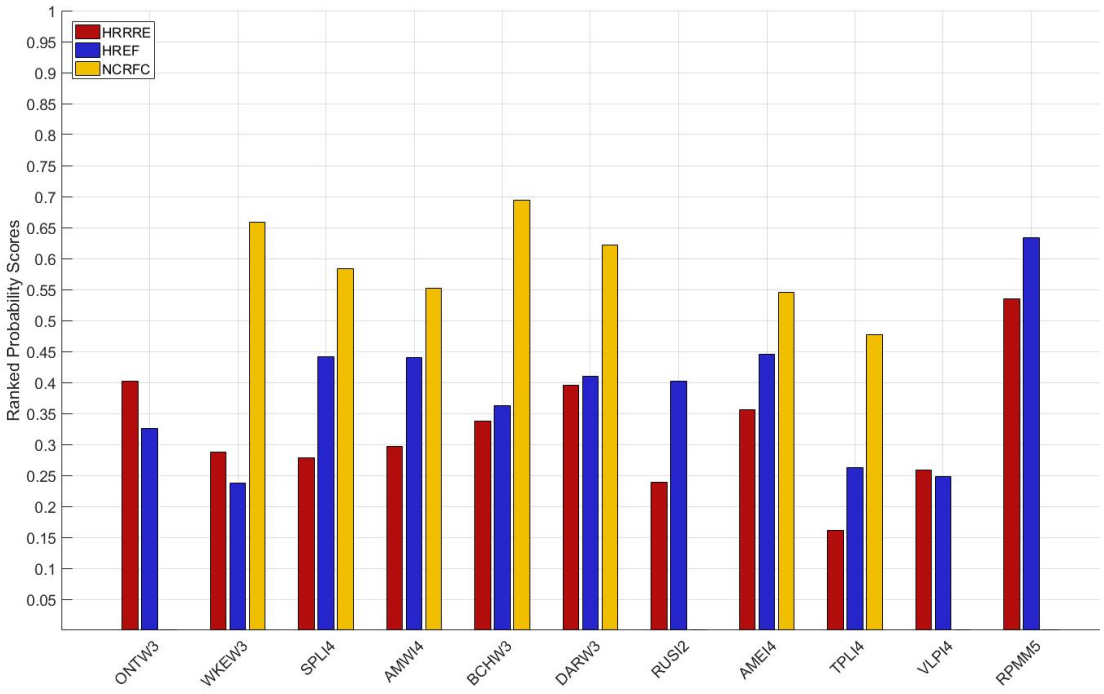

Figure 4: RPS for the eleven basins (labeled using NCRFC abbreviations) arranged in order from the smallest (left) to the largest (right) in area.

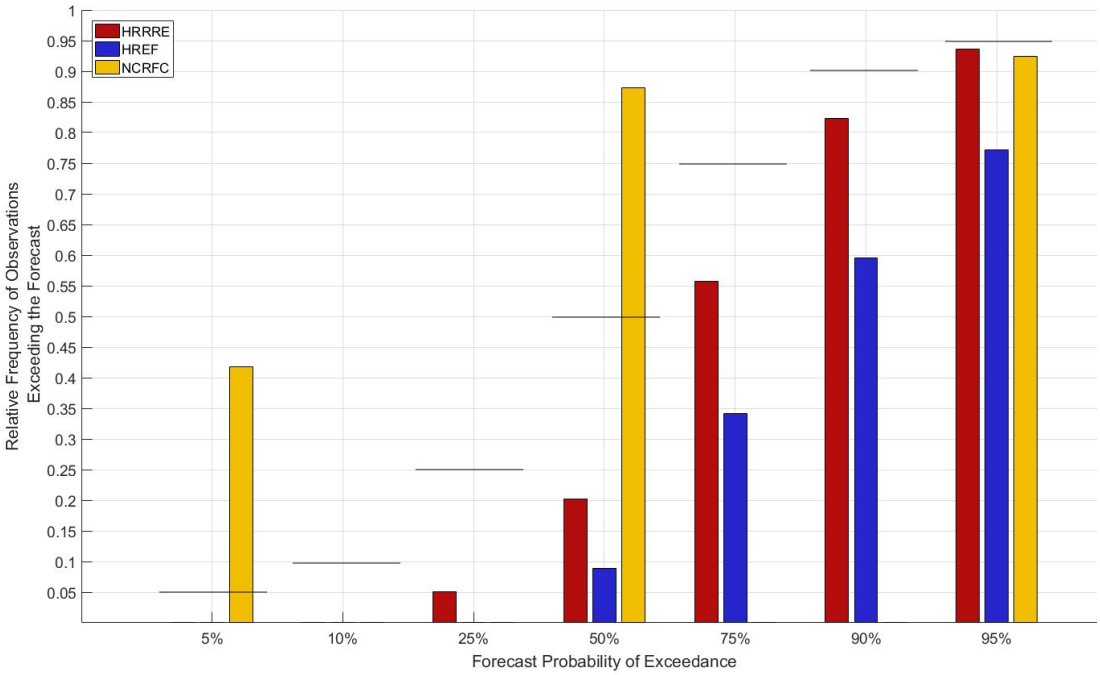

Figure 5: Relative frequency of times that observations exceeded the forecasts for different probability of exceedance values. All datasets use same 79 shared basins events.

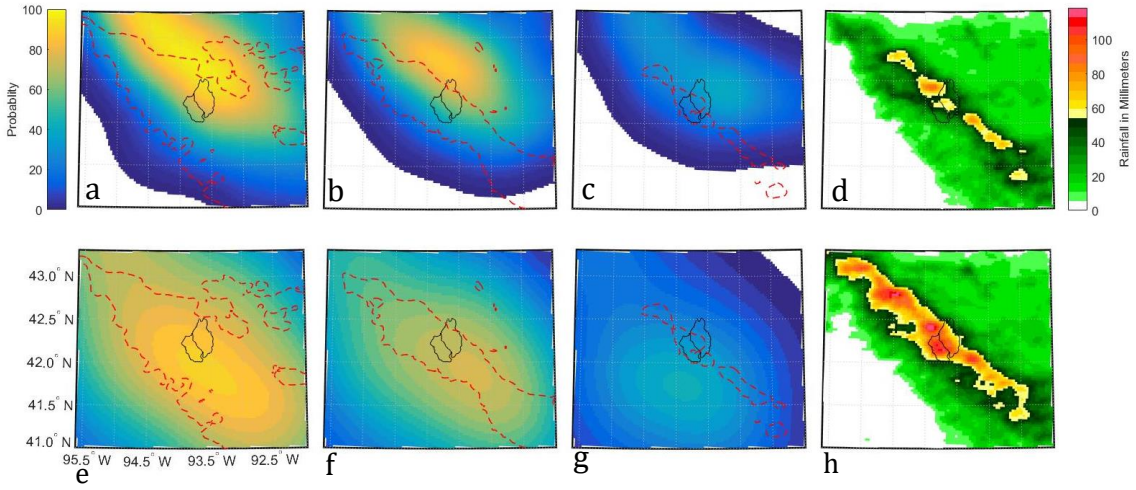

Figure 6: Show the probability values and rainfall accumulations over the central Iowa area for the June 14. The black outline represented the basins of Skunk (right) and Squaw Creek (left). Figure 6a-6g are generated during forecast hours 12-18, the peak accumulation time period. Figure 6a and 6e is the probabilities values at half an inch with the dash red line showing the measured half an inch from STAGE IV. Figure 6b and 6f shows the one-inch probabilities and measured precipitation. Figure 6c and 6g are associated with two inches. Figures 6a-6c are HRRRE while the Figures 6e-6g are HREF. Figure 6d is Stage IV measured precipitation for hours 12-18. Figure 6h is the total accumulation of rainfall over the 36 hours of the HRRRE and HREF forecasts.

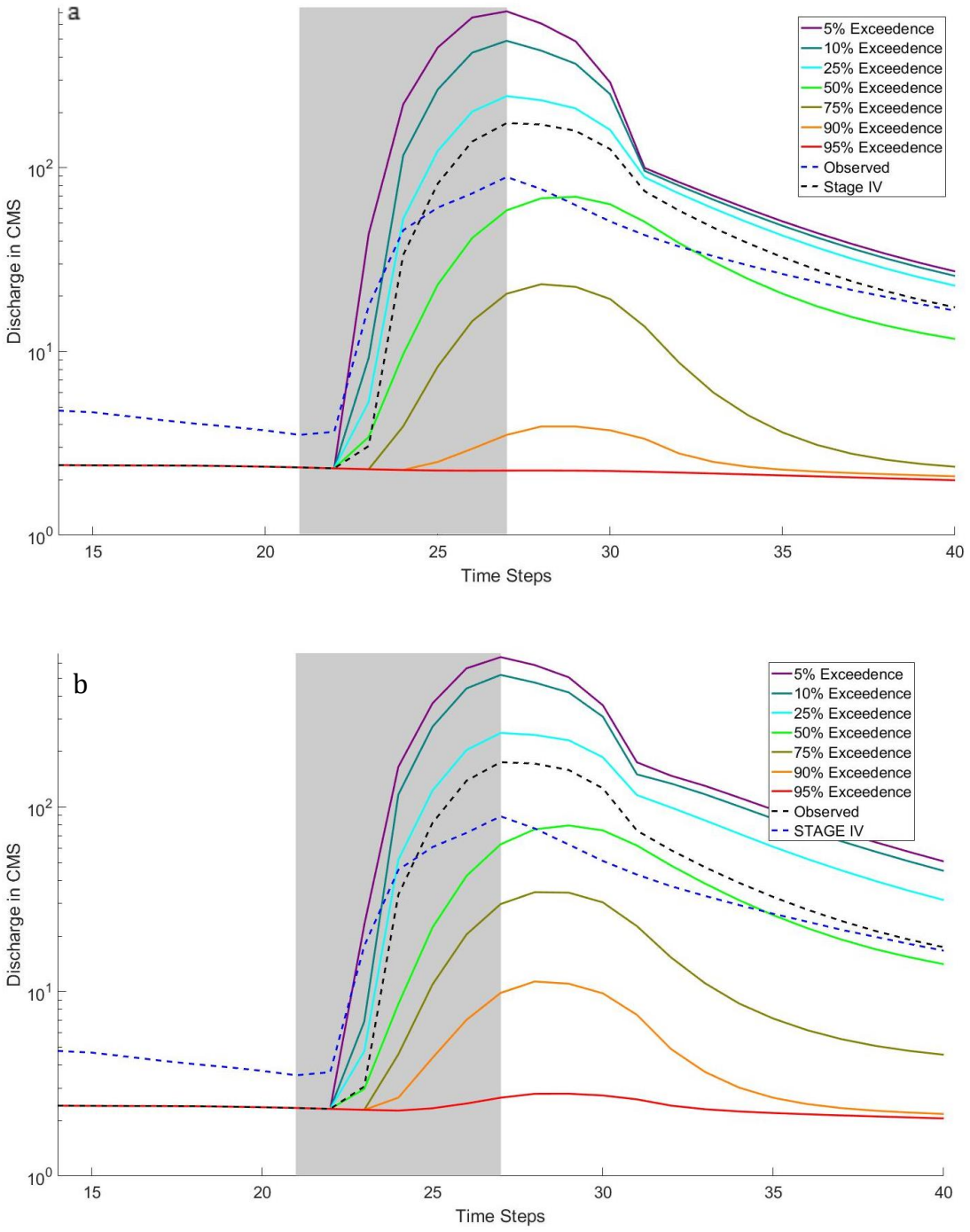

Figure 7: Hydrographs of the Skunk River for the June 14th flash flood event. The figures show the discharge for Skunk River at seven different probability values, the observed discharge from the USGS gauge and the discharge with the measured Stage IV data from NCRFC. The Figure 7a shows a hydrograph from the HRRRE probability values, and Figure 7b from the HREF values. The gray box in the background signifies the time steps where model values were inputted.

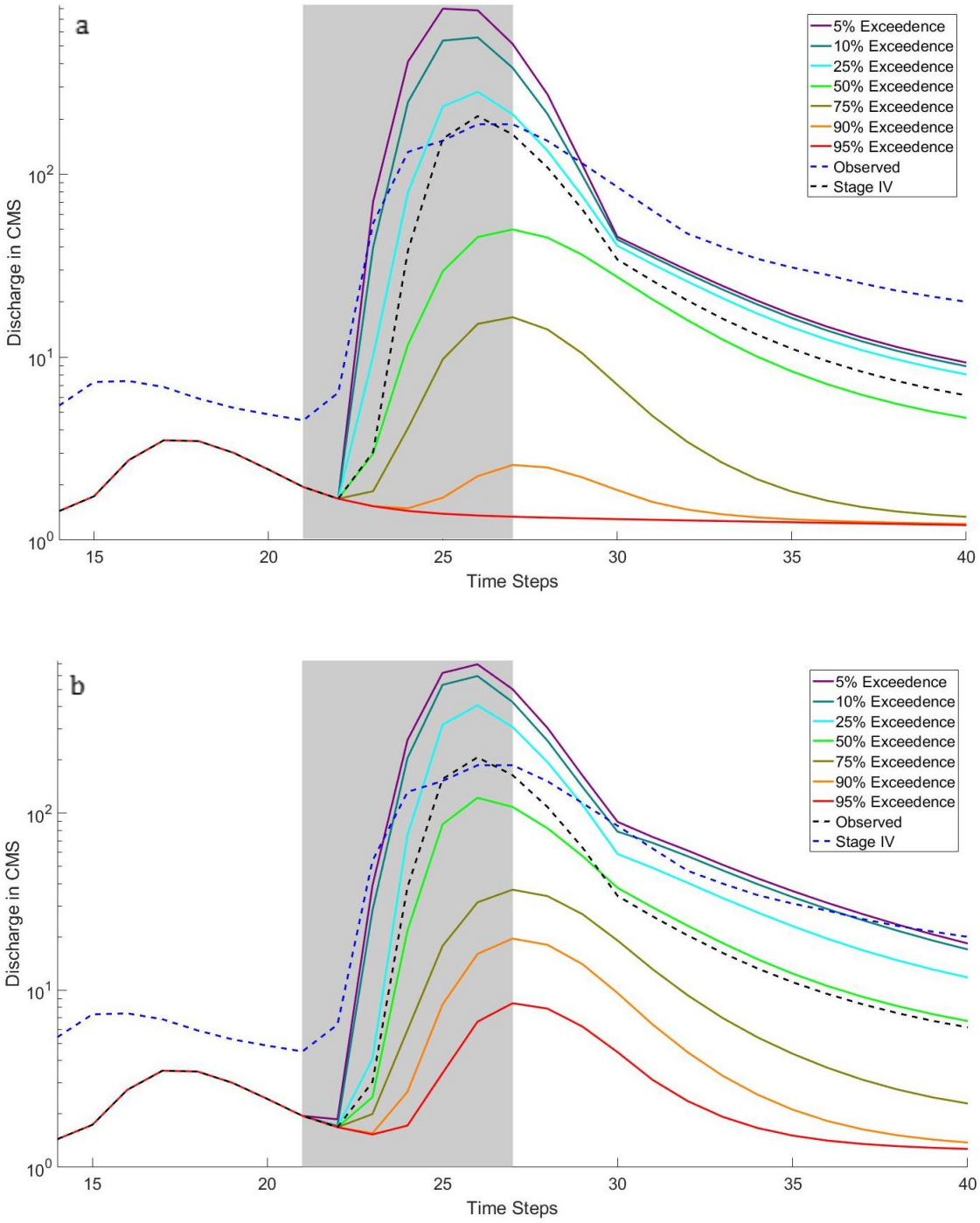

Figure 8: Same as in Figure 7, except for the Squaw Creek.

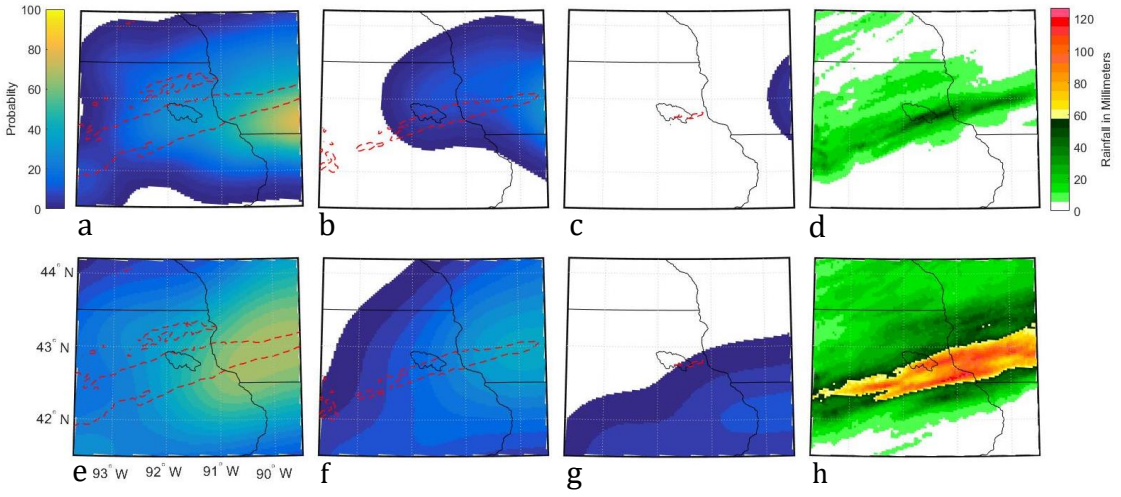

Figure 9: As in Figure 6, except for northeastern Iowa and the outline is the Volga River Basin for the September 30 - October 2 case, forecast hours 24-30.

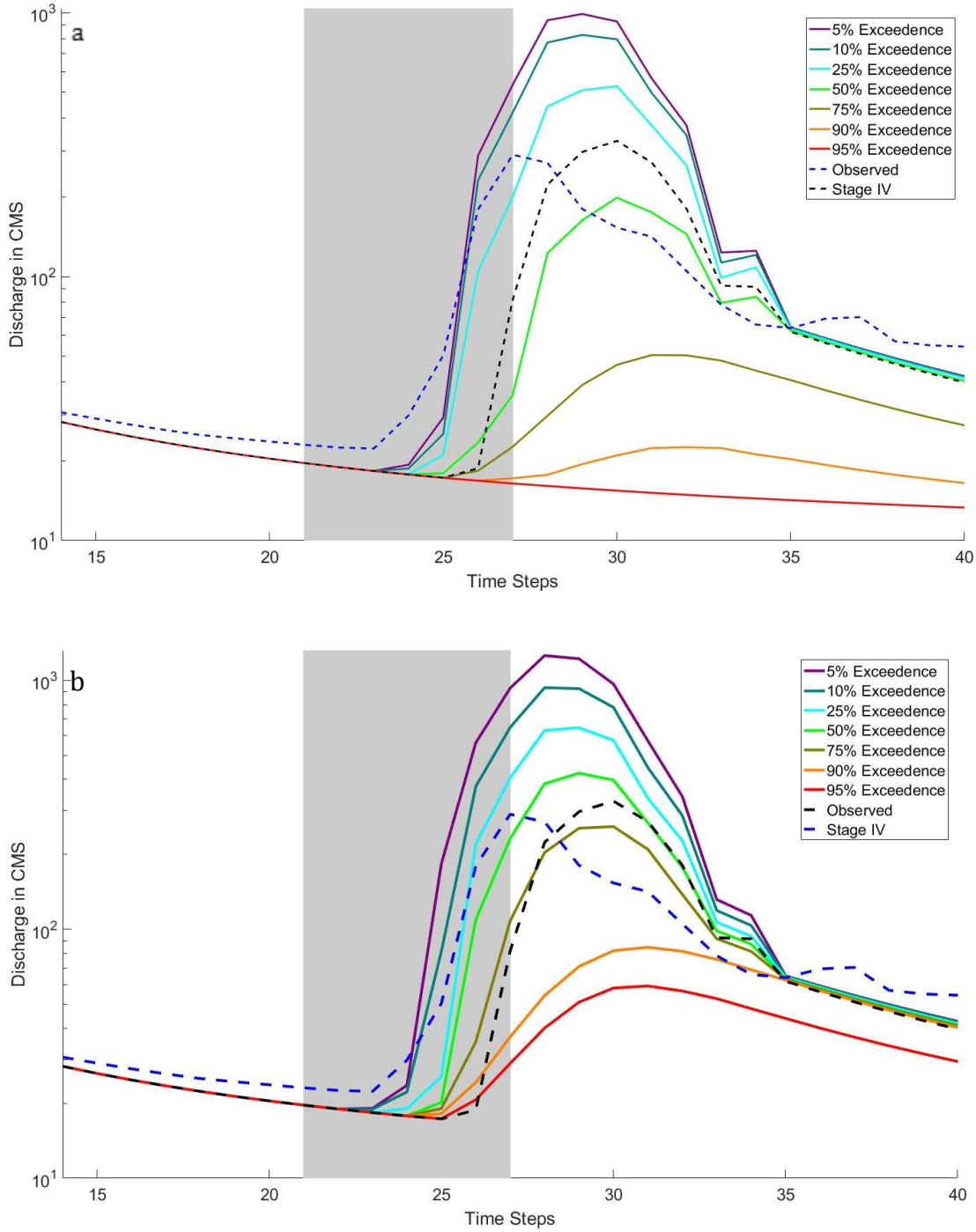

Figure 10: As in Figure 7, except for the Volga River on September 30-October 2.

CHAPTER 3 GENERAL CONCLUSIONS

The present study looked at an innovative technique of using probabilistic quantitative precipitation forecasts to predict streamflow changes in the near future. This technique was used as a possible method to mitigate two difficulties seen in current forecasting practices or discussed in past studies: Firstly, the technique could allow streamflow forecasts to be generated before the rainfall had started, thereby allowing emergency managers and the general public to be more prepared or evacuated before flooding could occur. Secondly, by using a precipitation product generated from an ensemble/probabilistic output, which limited problems seen in other studies, where errors in spatial or magnitude of accumulated precipitation constrained the forecasts produced by the hydrologic models (Davolio et al. 2008, Cuo et al. 2011, Wu et al. 2014, Seo et al. 2018).

The results show that both the HRRRE's and HREF's PQPFs were able to capture all streamflow changes from the 109 events within their probability values. Additionally, it was observed that predict streamflow changes generated from the predicted HRRRE PQPFs produced solutions that could be more similar to expected outcomes then systems currently used. This answer is most likely the result of the HRRRE model creating PQPF that was more accurate to observed rainfall.

Finally, the findings display that both the HRRRE and HREF PQPFs streamflow changes forecasts suffered from large biases at their low probability of exceedance forecasts. Future work is needed to examine possible techniques to mitigate these errors, whether it be done by calibration of probability-generated

rainfall amounts or spatially arranging the rainfall probabilities before calculating the basin averaged precipitation. Overall, the findings show that it is possible to generate streamflow forecasts that capture future peak discharge events consistently.

REFERENCES

Aligo, E., B. Ferrier, J. Carley, E. Rogers, M. Pyle, S.J. Weiss, and I.L. Jirak, 2014: Modified microphysics for use in high-resolution NAM forecasts. In *27th Conf. on Severe Local Storms*. Madison, WI, Amer. Meteor. Soc, https://ams.confex.com/ams/27SLS/videogateway.cgi/id/28414?recordingid=28414

Avolio E., S. Federico, M. Miglietta, T. Lo Feudo, C. Calidonna, and A. Sempreviva, 2017: Sensitivity analysis of WRF model PBL schemes in simulating boundary-layer variables in southern Italy: An experimental campaign. *Atmospheric Research,* **192,** 58-71, doi: 10.1016/j.atmosres.2017.04.003

Azizan I., K. S. A. Bin Abdul, and S. S. K. Raju, 2018: Fitting Rainfall Data by Using Cubic Spline Interpolation. *MATEC Web of Conferences*, 225, 05001.

Banks R., J. Tiana-Alsina, J. Baldasano, F. Rocadenbosch, A. Papayannis, S. Solomos, and C. Tzanis, 2016: Sensitivity of boundary-layer variables to PBL schemes in the WRF model based on surface meteorological observations, lidar, and radiosondes during the HygrA-CD campaign. *Atmospheric Research*, **176-177**, 185-201. doi: 10.1016/j.atmosres.2016.02.024

Buizza, R., M. Miller, and T. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Quarterly Journal of The Royal Meteorological Society*, **125(560)**, 2887-2908.

Burnash, R. J., R. L. Ferral, and R. A. McGuire, 1973: A Generalized Streamflow Simulation System: Conceptual Modeling for Digital Computers. National Weather Service and State of California Department of Water Resources, 204 pp.

Carlberg B., W. Gallus, and K. Franz, 2018: Improving convective mode and streamflow forecasts through the use of convection-allowing ensembles, Iowa State University, 45-69

Cintineo, R., J. A. Otkin, M. Xue, and K. Fanyou, 2014: Evaluating the performance of planetary boundary layer and cloud microphysical parameterization schemes in convectionpermitting ensemble forecasts using synthetic GOES-13 satellite observations*. Mon. Wea. Rev*., **142**, 163–182. doi: https:// doi.org/10.1175/MWR-D-13-00143.1.

Clark, A., 2017: Generation of ensemble mean precipitation forecasts from convection-allowing ensembles*. Wea. Forecasting*, **32**, 1569–1583. doi: https://doi.org/10.1175/ WAF-D-16-0199.1.

Cuo, L., T.C. Pagano, and Q.J. Wang, 2011: A Review of Quantitative Precipitation Forecasts and Their Use in Short- to Medium-Range Streamflow Forecasting. *J. Hydrometeor*., **12**, 713–728. doi: https://doi.org/10.1175/2011JHM1347.1

Davolio, S., M. M. Miglietta, T. Diomede, C. Marsigli, A. Morgillo, and A. Moscatello, 2008: A meteo-hydrological prediction system based on a multi-model approach for precipitation forecasting. *Natural Hazards and Earth System Sciences*, **8(1)**, 143-159. doi: 10.5194/nhess-8-143-2008

Dowell, D., C. Alexander, T. Alcott, and T. Ladwig, 2018: HRRR Ensemble (HRRRE) guidance 2018 HWT spring experiment. Accessed 1 September 2018, https://rapidrefresh.noaa.gov/internal/pdfs/2018_Spring_Experiment_HRRRE_Docume ntation.pdf

Duda, J., and W. Gallus, 2013: The Impact of Large-Scale Forcing on Skill of Simulated Convective Initiation and Upscale Evolution with Convection-Allowing Grid Spacings in the WRF. *Weather And Forecasting,* **28 no. 4**, 994-1018. doi: 10.1175/WAF-D-13-00005.1

Dziubanski, D.J., Franz, K.J., 2016. Assimilation of AMSR-E snow water equivalent data in a spatially-lumped snow model. J. Hydrol. **540**, 26–39. http://dx.doi.org/10. 1016/j.jhydrol.2016.05.046

Ebert, E., 2001: Ability of a Poor Man's Ensemble to Predict the Probability and Distribution of Precipitation. *Monthly Weather Review,* **129**, 2461–2480. doi: 10.1175/1520-0493

_____, U. Damrath, W. Wergen, and M. Baldwin, 2003: THE WGNE ASSESSMENT OF SHORT-TERM QUANTITATIVE PRECIPITATION FORECASTS. *Bulletin of the American Meteorological Society*, **84(4)**, 481-492. doi: 10.1175/BAMS-84-4-Ebert

____, 2009: Neighborhood verification: A strategy for rewarding close forecasts. Wea. Forecasting, **24**, 1498– 1510. doi: https://doi.org/10.1175/2009WAF2222251.1.

Eckel, F. A., and C. F. Mass, 2005: Aspects of effective mesoscale, short-range ensemble forecasting. *Wea. Forecasting*, **20**, 328–350. doi: 10.1175/WAF843.1

Franz, K. J., H. C. Hartmann, S. Sorooshian, and R. Bales, 2003: Verification of National Weather Service ensemble streamflow predictions for water supply forecasting in the Colorado River basin. *J. Hydrometeor*., **4**, 1105–1118, doi: 10.1175/1525- 7541

Fritsch, J. M., and R. E. Carbone, 2004: Improving quantitative precipitation forecasts in the warm season: A USWRP research and development strategy. *Bull. Amer. Meteor. Soc.*, **85**, 955–965. dio: 10.1175/BAMS-85-7-955

Georgakakos, K., and M. Hudlow, 1984: Quantitative Precipitation Forecast Techniques for Use in Hydrologic Forecasting. *Bulletin of the American Meteorological Society, **65*(11),** 1186-1200. doi: https://doi.org/10.1175/1520-0477(1984)065%3C1186:QPFTFU%3E2.0.CO;2

Gilleland, E., D. Ahijevych, B. G. Brown, B. Casati, and E. E. Ebert, 2009: Intercomparison of spatial forecast verification methods. Wea. Forecasting, **24**, 1416–1430, https://doi.org/ 10.1175/2009WAF2222269.1

Golding, B.W., 2000: Quantitative Precipitation Forecasting in the UK. *Journal of Hydrology,* **239 no. 1**, 286-305. doi: 10.1016/S0022-1694(00)00354-1

Hong, S.Y. and J.O.J Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc*, **42(2)**, 129-151. http://www2.mmm.ucar.edu/wrf/users/phys_refs/MICRO_PHYS/WSM6.pdf

Hong, S.–Y., Y. Noh, J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev*., **134**, 2318–2341. doi:10.1175/MWR3199.1

Hou, D., M. Charles, Y. Luo, Z. Toth, Y. Zhu, R. Krzysztofowicz, Y. Lin, P. Xie, D. Seo, M. Pena, and B. Cui, 2014: Climatology-Calibrated Precipitation Analysis at Fine Scales: Statistical Adjustment of Stage IV toward CPC Gauge-Based Analysis. *J. Hydrometeor.*, **15**, 2542–2557. doi: https://doi.org/10.1175/JHM-D-11-0140.1

Janjic, Z., 1994: The Step–Mountain Eta Coordinate Model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. Mon. Wea. Rev., **122**, 927–945. doi: 10.1175/1520-0493

Kitchen M., and P. M. Jackson, 1993: Weather radar performance at long range—Simulated and observed. *J. Appl. Meteor*., **32**, 975–985, doi: 10.1175/1520-0450(1993)032

Krajewski, W. F., and Coauthors, 2017: Real-time flood forecasting and information system for the state of Iowa. *Bull. Amer. Meteor. Soc*., **98**, 539–554, doi: 10.1175/BAMS-D-15- 00243.1.

Leith, C. E., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev*., **102**, 409– 418, doi: 10.1175/1520-0493

Leutbecher, M., T. Palmer, 2008: Ensemble forecasting. *Journal of Computational Physics,* **227**, 3515–3539. doi: 10.1016/j.jcp.2007.02.014

Moser, B., W. Gallus, Jr, & R. Mantilla, 2015: An Initial Assessment of Radar Data Assimilation on Warm Season Rainfall Forecasts for Use in Hydrologic Models, *Geological and Atmospheric Sciences Publications,* **30**, 1491-1520, doi: 10.1175/WAF-D-14-00125.1

Nakanishi, M., and H. Niino, 2009: Development of an improved turbulence closure model for the atmospheric boundary layer. *J. Meteor. Soc*. Japan, **87**, 895–912. doi:10.2151/jmsj.87.895

Nelson, B. R., O. P. Prat, D. J. Seo, and E. Habib, 2016: Assessment and implications of NCEP Stage IV quantitative precipitation estimates for product intercomparisons. *Wea. Forecasting*, **31**, 371–394, doi: 10.1175/WAF-D-14-00112.1

Nielsen, E. and R. Schumacher, 2016: Using Convection-Allowing Ensembles to Understand the Predictability of an Extreme Rainfall Event. *Monthly Weather Review* 144, **no. 10**, 3651-3676, doi:10.1175/MWR-D-16-0083.1

NOAA (National Centers for Environmental Information), Storm Events Database. Accessed 28 January 2019, https://www.ncdc.noaa.gov/stormevents/

_____, (National Centers for Environmental Information), U.S. Billion-Dollar Weather and Climate Disasters 2017. Accessed 24 March 2019, https://www.ncdc.noaa.gov/billions/events/US/2017

NWS, 2018: Summary of natural hazard statistics for 2017 in the United States. Accessed 28 January 2019, http://www.nws.noaa.gov/om/hazstats/sum17.pdf.

Nguyen, P., A. Thorstensen, S. Sorooshian, K. Hsu, and A. AghaKouchak, 2015: Flood forecasting and inundation mapping using HiResFlood-UCI and near-real-time satellite precipitation data: The 2008 Iowa flood. *J. Hydrometeor*., **16**, 1171–1183, doi: 10.1175/JHM-D-14-0212.1.

Rabuffetti D. and S. Barbero, 2005: Operational hydro-meteorological warning and real-time flood forecasting: the Piemonte Region case study. *Hydrology and Earth System Sciences*, **9(4)**, pp.457–466. doi: 10.5194/hess-9-457-2005

Roberts, B., I. Jirak, A. Clark, S. Weiss, and J. Kain, 2018: Post-processing and visualization techniques for convection-allowing ensembles. *Bulletin of the American Meteorological Society,* D-18-0041.1. doi: http://dx.doi.org/10.1175/BAMS-D-18-0041.1

Schaffer C., W. Gallus, and M. Segal, 2011: Improving probabilistic ensemble forecasts of convection through the application of QPF–POP relationships. *Wea. Forecasting*, **26**, 319–336, doi: 10.1175/2010WAF2222447.1

Seo, B.-C., F. Quintero, and W.F. Krajewski, 2018: High-Resolution QPF Uncertainty and Its Implications for Flood Prediction: A Case Study for the Eastern Iowa Flood of 2016. *Journal of Hydrometeorology*, **19(8)**, pp.1289–1304. doi: 10.1175/JHM-D-18-0046.1

Singh, V., 2018: Hydrologic modeling: progress and future directions. *Geoscience Letters*, **5(1)**, pp.1–18. Doi: 10.1186/s40562-018-0113-z

United States Geological Survey, 2018: South Skunk River near Ames, IA, USGS, October 2018, https://waterdata.usgs.gov/ia/nwis/uv?site_no=05470000

_____, 2018: Squaw Creek at Ames, IA, USGS, October 2018, https://waterdata.usgs.gov/ia/nwis/uv/?site_no=05470500

_____, 2018: Volga River at Littleport, IA, USGS October 2018, https://waterdata.usgs.gov/usa/nwis/uv?site_no=05412400

University Corporation for Atmospheric Research, 2018: Image Archive Meteorological case study selection kit, October 2018, http://www2.mmm.ucar.edu/imagearchive/

Thompson, G., and T. Eidhammer, 2014: A study of aerosol impacts on clouds and precipitation development in a large winter cyclone. J. Atmos. Sci., **71.10**, 3636-3658, doi: 10.1175/JAS-D-13-0305.1

Trevor, A. 2018: HRRRE Setup Conditions.

Villarini, G., M. Pradeep, K. Witold, and M. Robert, 2008: Rainfall and Sampling Uncertainties: A Rain Gauge Perspective*. Journal of Geophysical Research,* **113(D11)**, doi: 10.1029/2007JD009214

Wilks, D.S., 2011: *Statistical methods in the atmospheric sciences*. *3rd ed.,* Oxford, Waltham, MA: Academic Press.

Wilson, J.W. and E.A. Brandes, 1979: Radar Measurement of Rainfall—A Summary. Bull. *Amer. Meteor. Soc*., **60**, 1048–1060, doi: 10.1175/1520-0477(1979)060

Wu, J., G. Lu, and Z. Wu, 2014: Flood Forecasts Based on Multi-model Ensemble Precipitation Forecasting Using a Coupled Atmospheric-hydrological Modeling System. *Natural Hazards,* **74 no. 2**, 325-40, dio: 10.1007/s11069-014-1204-6